

Print out

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

OM protein - protein search, using sw model

Run on: July 6, 2006, 01:28:04 ; Search time 316 Seconds

Sequence: 1 MTNCKYKURKLSQLVGLSVGTT.....PTIIGAGILSKRRDTEGN 1046

Title: US-10-771-931-1  
Perfect score: 5357  
Scoring table: BLOSUM62  
GapOp 10.0 , Gapext 0.5

Searched: 2849598 seqs, 925015592 residues

Total number of hits satisfying chosen parameters: 2849598

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt 7.2;\*  
1: uniprot\_sprot;\*  
2: uniprot\_trembl;\*

**pred.** No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	5357.5	100.0	1046	084941_STREY
2	3575.5	66.7	1027	093T51_STREY
3	3500.5	65.4	1026	094ZP4_STREY
4	3412.5	63.7	1023	093T53_STREY
5	3371.5	62.9	1025	094507_STREY
6	3370.0	62.9	1026	092836_STREY
7	3357.5	62.7	1025	092532_STREY
8	3164.5	59.1	1013	09L953_STREY
9	3132.5	58.5	1015	093T50_STREY
10	2973.5	55.5	1013	093T52_STREY
11	2878.5	53.7	954	06YBUG_STREY
12	2874.0	52.9	1022	093T54_STREY
13	2779.5	51.9	963	093T49_STREY
14	2654.5	49.6	1029	09X2V2_STREY
15	2571.0	48.0	862	09RPA2_STREY
16	2531.0	47.2	923	095S32_STREY
17	2488.0	46.4	865	095S49_STREY
18	2485.0	46.0	873	095S40_STREY
19	2460.5	45.9	872	095S43_STREY
20	2433.5	45.5	854	095S38_STREY
21	2329.0	43.5	881	09RQ07_STREY
22	2138.5	39.9	1091	006556_STREY
23	2102.0	39.2	746	08VS95_STREY
24	1458.5	27.2	409	09X899_STREY
25	1391.5	26.0	456	09XCK8_STREY
26	1388.5	25.9	471	09EZB18_STREY
27	1376.5	25.7	447	09RQ07_STREY
28	1343.5	25.1	452	09RQ02_STREY
29	1331.5	24.9	439	09RQ08_STREY
30	1276.0	23.8	454	09S418_STREY
31	1262.0	23.6	423	09EZB16_STREY

SG	RESULT 1	084941_STREY	PRELIMINARY	PRT: 1046 AA.
ID	084941_STREY			
AC	084941;1			
DT	01-Nov-1998, Integrated into UniProtKB/TREMBL.			
DT	01-Nov-1998, Sequence version 1.			
DT	07-FEB-2006, entry version 25.			
DR	Serum opacity factor.			
DR	Name=sbof;			
OS	Streptococcus pyogenes.			
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;			
OC	Streptococcus.			
OX	NCBI_TaxID=114;			
RN	[1]			
RC	NUCLEOTIDE SEQUENCE.			
RX	MEDLINE:99233508; PubMed:10316862;			
RA	Courtney H.S.; Hasty D.L.; Li Y.; Chiang H.C.; Thacker J.L.; Dale J.B.;			
RA	"Serum opacity factor is a major fibronectin-binding protein and a virulence determinant of M type 2 Streptococcus pyogenes"; Mol. Microbiol. 32:89-98 (1999).			
CC	Copyrighted by the UniProt Consortium, see <a href="http://www.uniprot.org/terms">http://www.uniprot.org/terms</a>			
CC	Distributed under the Creative Commons Attribution-NonDerivs License			
CC				
EMBL	AT019830; AAC32396.1; -; Genomic_DNA.			
DR	GO: GO-0009866; Cicell surface; IEA.			
DR	GO: GO-0009275; Cicell wall (semu Gram-positive Bacteria); IEA.			
DR	GO; GO-00016020; C:membrane; IEA.			
DR	GO; GO-0007155; P:cell adhesion; IEA.			
DR	InterPro; IPR004237; Fibron repeat_Bd.			
DR	InterPro; IPR005877; Gpos_YSIRK.			
DR	InterPro; IPR001999; Gram_pos_anchor.			
DR	InterPro; IPR002035; VWF_A.			
DR	Pfam; PP02986; Fn_bind; 3.			
DR	Pfam; PP00746; Gram_pos_anchor; 1.			
DR	Pfam; PP00052; VWA; 1.			
DR	SMART; SM00327; VWA; 1.			
DR	TIGRFAMS; TIGR01168; YSIRK_signal; 1.			
DR	PROSITE; PS08471; GRAM_POS_ANCHORING; 1.			
DR	PROSITE; PS02347; VWF_A; 1.			
SEQUENCE	1046 AA; 114914 MW; B12CP6B88059B62 CRC64;			

Qy	Query	Match	100.0%; Score 5357; DB 2; Length 1046; Matches 1046; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	DB	Best Local Similarity 100.0%; Pred. No. 3.1e-205;	
Qy	Qy	1 MTNCKYKURKLSQLVGLSVGTT.....PTIIGAGILSKRRDTEGN 60	
Db	Db	1 MTNCKYKURKLSQLVGLSVGTT.....PTIIGAGILSKRRDTEGN 60	
Qy	Qy	61 GAVTTATTATGGPOSTAVASPTPOAQIAWAASTSSASSSDGKAVT 120	

Db	61	GAATTTATTATGGPOSTPAAATPOPOQIAPVAATSSASSSDGKAPAVT	120	DT
Qy	121	SSTSFRSTPAASSNSNQSAARTPQTOMEVKTYTDKENSKLNDKGKPRTSNSYNE	180	DE Serum opacity factor; VT21.
Db	121	SSTSFRSTPAASSNSNQSAARTPQTOMEVKTYTDKENSKLNDKGKPRTSNSYNE	180	OS Streptococcus pyogenes.
Qy	181	KOTKLLIRNRDGKLDIVDVTVKNEGTIDVTVKPQIDEGADVMALDVSKRMSE	240	OC Bacteria; Firmicutes; Lactobacillales; Streptococaceae;
Db	181	KOTKLLIRNRDGKLDIVDVTVKNEGTIDVTVKPQIDEGADVMALDVSKRMSE	240	OC Streptococcus.
Qy	241	DDFRNARAKIKKGVKLTTSKASASNSDNDEHKNSNSVRMLTYFVIBSNPIDISCTEQ	300	NCBI_TaxID=114;
Db	241	DDFRNARAKIKKGVKLTTSKASASNSDNDEHKNSNSVRMLTYFVIBSNPIDISCTEQ	300	RN [1]
Qy	301	LDKLUDLRLRKAKANPWTGVDLQGAIKAREIFNGKEKEKGKRRHIVFSQGSTPSYE	360	RP NUCLEOTIDE SEQUENCE.
Db	301	LDKLUDLRLRKAKANPWTGVDLQGAIKAREIFNGKEKEKGKRRHIVFSQGSTPSYE	360	RX MEDLINE=2164287; PubMed=1782509;
Qy	361	LQNSVREDKTKLSRSLGAVTSSNPFLPPWPPFLPNHFKNIMDDVNLVKGQGIGL	420	RA Gilien C.M., Towers R.J., McMillan D.J., Delvecchio A.,
Db	361	LQNSVREDKTKLSRSLGAVTSSNPFLPPWPPFLPNHFKNIMDDVNLVKGQGIGL	420	RA Sriprakash K.S., Currie B., Kreikemeyer B., Chhatwal G.S.,
Qy	421	DNQSTLTLISTSSSLAGAFLGGGSILTELTKESTKGDKIENPPDYTRVGECHFHSP	480	RA Walker M.J.,
Db	421	DNQSTLTLISTSSSLAGAFLGGGSILTELTKESTKGDKIENPPDYTRVGECHFHSP	480	RT RT the Northern Territory of Australia against <i>Streptococcus pyogenes</i>
Qy	481	SEKKTGTEPPKSLIPKPKELFNNKNUODKSMTWIDKSLTERIQAKOETLMKUL	540	RT serum opacity factor.;
Db	481	SEKKTGTEPPKSLIPKPKELFNNKNUODKSMTWIDKSLTERIQAKOETLMKUL	540	RL Microbiology 148:169-178 (2002).
Qy	541	EYLFYKREHYMMNLATSAAKQMQEGITFYSDVDTLKTTSRKVQVESTEDKCKE	600	CC Copyrighted by the UniProt Consortium, see <a href="http://www.uniprot.org/terms">http://www.uniprot.org/terms</a>
Db	541	EYLFYKREHYMMNLATSAAKQMQEGITFYSDVDTLKTTSRKVQVESTEDKCKE	600	CC Distributed under the Creative Commons Attribution Non-Commercial License
Qy	601	KOREDEIEKERNEKEKDNYLKMSECGKDPFDVKEKEPKDQLINTVTEFEDGVNKON	660	DR EMBL; ARF367014; AAAX52968.1; -; Genomic_DNA.
Db	601	KOREDEIEKERNEKEKDNYLKMSECGKDPFDVKEKEPKDQLINTVTEFEDGVNKON	660	DR GO; GO:0009275; C:cell wall (sensu Gram-positive Bacteria); IEA.
Qy	661	SKQVSSENNLSNLYNKSYTHKAASDASWLSLSKESIATWTSKQKLAFAEKNSSLTP	720	DR GO; GO:0016020; C:membrane; IEA.
Db	661	SKQVSSENNLSNLYNKSYTHKAASDASWLSLSKESIATWTSKQKLAFAEKNSSLTP	720	DR GO; GO:0007155; P:cell adhesion; IEA.
Qy	721	KYKLUQVNQKQKLNQKRNTRDTSSENKTSVTKOIISNTVYKINOBVKGKLNQDVKLT	780	DR InterPro; IPRO04237; Fibron repeat_bd.
Db	721	KYKLUQVNQKQKLNQKRNTRDTSSENKTSVTKOIISNTVYKINOBVKGKLNQDVKLT	780	DR InterPro; IPRO01899; Gpos_YSKR.
Qy	781	YKTCETVPDPVEGVYPIPKPUPEMPTLPATNPYPPDIPPQLPQDDELIISGGHG	840	DR InterPro; IPRO02035; VWF_A.
Db	781	YKTCETVPDPVEGVYPIPKPUPEMPTLPATNPYPPDIPPQLPQDDELIISGGHG	840	DR Pfam; PF03986; Pf_b1nd; 3.
Qy	841	PSDIVDTGTGARGGAQNGVUSTQENRDPIDEDTDTGPMSSNSDATVEEDTAKRP	900	DR Pfam; PF00746; Gram_Pos_anchor; 1.
Db	841	PSDIVDTGTGARGGAQNGVUSTQENRDPIDEDTDTGPMSSNSDATVEEDTAKRP	900	DR SMART; SM0327; VWF_1.
Qy	901	DVLYGQSDPIDIETDTPSVGSSNDATVEEDTVPKRDPDSVGSQSDPIDIETDTPGM	960	DR TIGR01168; YSTRK_signal; 1.
Db	901	DVLYGQSDPIDIETDTPSVGSSNDATVEEDTVPKRDPDSVGSQSDPIDIETDTPGM	960	DR PROSITE; PS50847; GRAM_Pos_ANCHORING; 1.
Qy	961	SGSGNATIEEDTRKRVTHDFNQPQKEPKPNEQESLSLSPQAPTYKAHLPLASGDKREA	1020	DR PROSITE; PS5034; VWF_A; 1.
Db	961	SGSGNATIEEDTRKRVTHDFNQPQKEPKPNEQESLSLSPQAPTYKAHLPLASGDKREA	1020	DR Sequence; 1027_AA; 112542_MW; 0482398598RA4C7A_CRC64;
Qy	1021	SFTIAAPTTGAAGLISKCRDTEGN 1046	1020	Query Match 66.7%; Score 3755.5; DB 2; Length 1027; Best Local Similarity 70.1%; Pred. No. 2.8e-134; Matches 751; Conservative 98; Mismatches 153; Indels 69; Gaps 26; DR .
Db	1021	SFTIAAPTTGAAGLISKCRDTEGN 1046	1020	DR .
Qy	1021	SFTIAAPTTGAAGLISKCRDTEGN 1046	1020	DR .
Db	1021	SFTIAAPTTGAAGLISKCRDTEGN 1046	1020	DR .

Printed ②

Page 1

GenCore version 5.1.9  
Copyright (c) 1993 - 2006 Biocceleration Ltd.

On protein - protein search, using sw model  
Run on: July 6, 2006, 01:31:29 ; Search time 48 Seconds  
(without alignments)  
2096.724 Million cell updates/sec

Title: US-10-771-931-1  
Perfect score: 5357  
Sequence: 1 MTNCKYKRLKLGSVLSVGT.....PTIIGAAGLISKKRDRDTGCG 1046  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext: 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries:

Database :

1: PIR:80;\*  
2: pir1;\*  
2: pir2;\*  
3: pir3;\*  
4: pir4;\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

### SUMMARIES

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3346.5	62.5	1025	2	\$69790		fibronectin-binding protein II precursor - Streptococcus pyogenes
2	2138.5	39.9	1091	2	\$33850		fibronectin-binding
3	409.5	7.6	659	2	\$4043		adhesin - Streptococcus
4	394	7.4	638	2	\$54418		fibronectin-binding
5	354	6.6	961	2	\$9053		hypothetical protein
6	351.5	6.6	1018	2	A32192		fibronectin-binding
7	339	6.3	1038	2	H9053		hypothetical protein
8	338.5	6.3	940	2	S19702		fibronectin-binding
9	325.5	6.1	1039	2	T30856		protein F2 - Streptococcus
10	306	5.7	1117	2	S33851		fibronectin-binding
11	302	5.6	1164	1	RCSBQAC		IGA Fc receptor protein
12	289	5.4	1134	2	A60234		IGA Fc receptor protein
13	282	5.3	1092	2	S42798		fibronectin-binding
14	268.5	5.0	1875	2	S38173		myosin-like protein
15	263.5	4.9	1726	1	S420GM		major merozoite surface protein
16	263.5	4.9	1726	2	A45948		major merozoite surface protein
17	263	4.9	1726	2	T18296		myosin heavy chain
18	257.5	4.8	2481	2	D90011		FmtB protein limpoprotein
19	248.5	4.6	2722	2	T20532		hypothetical protein
20	248	4.6	2663	1	S28261		centromere protein
21	247	4.6	1526	2	A45605		mature-parasitein
22	245	4.6	1086	2	S1752		major merozoite surface protein
23	241.5	4.5	853	2	T1505		hypothetical protein
24	241	4.5	2269	2	T38677		aggregating protein
25	240.5	4.5	2262	2	S25624		hypothetical protein
26	240	4.5	1127	2	T28317		ORF MSV16 hypothetical protein
27	240	4.5	1141	2	E8824		hypothetical protein
28	239.5	4.5	1125	2	B90598		membrane nuclelease
29	239.5	4.5	2508	2	S61441		surface-associated

### ALIGNMENTS

#### RESULT 1

569790  
Species: Streptococcus pyogenes  
C.I.Date: 14-Feb-1997 #Sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C.Accession: S69790; S72665

R.Kreikemeyer, B.; Talay, S.R.; Chhatwal, G.S.  
Mol. Microbiol. 17, 137-145, 1995  
A.Title: Characterization of a novel fibronectin-binding surface protein in group A streptococci

A.Reference number: S69790; MUID:96020668; PMID:7476200

A.Accession: S69790

A.Status: nucleic acid sequence not shown

A.Molecule type: DNA

A.Residues: 1-1025 <KRE>

A.External source: strain A75

A.Note: the authors translated the initiation codon TGG for residue 1 as Leu

R.Kreikemeyer, B.  
Submitted to the EMBL Data Library, December 1994

A;Reference number: S72666

A;Molecule type: DNA

A;Residues: 1-19, 'T', 21-145, 'T', 147-1025 <KRW>

A;Cross-references: UNIPARC:UPI0000AF5BB; EMBL:X83303; NID:91070387; PIDN:CA58282.1;

C;Genetics: S72666

A;Gene: S72666

A;Start codon: TGG

C;Keywords: fibronectin binding; membrane bound

F;1-347/domain: signal sequence #status predicted <SG>

F;35-1025/Product: fibronectin-binding protein II #status predicted <MAT>

F;848-860/Region: fibronectin binding

F;869-907/Region: fibronectin binding

F;908-946/Region: fibronectin binding

F;930-994/Region: membrane anchor cleavage motif

F;999-1016/Domain: transmembrane #status predicted <TM>

Query Match 62.5%; Score 3346.5; DB 2; Length 1025;

Best Local Similarity 64.0%; Pred. No. 3.1e-130; Mismatches 181; Indels 61; Gaps 19;

Matches 682; Conservative 181; Insertions 181; Deletions 61; Gaps 19;

Query 1 MTNCKYKRLKLGSVLSVGT.....PTIIGAAGLISKKRDRDTGCG 52

Db 61 TAAS-GASSEA-----TATANGCPQSPATEPATPQQAPAPASPTVSSSD 113

QY 111 SDGKPAQVTSSTSPTAAASSNSQASAEATPEOTVEKTYDVKNSKNUKGRT 170

QY 114 SDAKTPKA-ASTISSATTASPSNSKRNAAEPAPOMDNEQYKIKDENSITVAD-- 169

protein F07All\_6 ( protein involved in aggregation protein fibrinogen-binding hypothetical prote probable calcium-b fibroprotein - hypothetical protein - hypothetical prote microtubule bindin muramidase-release salivary agglutini major merozoite su

Qy 171 PKIGSSVNEKOTKLIRURDGSKLARDIVUTVTRTKNTEDTIVTVTPKQIDEGADWMA 230  
 Db 170 -----KAKOLKIRDNPKDLPVKRVEKGDTGLTUKMPQIDEGADWA 221  
 Qy 231 LLDYNSKNSBDDTNAKOKUKVLTJTSKASNSDDEHKNSRNSTMLMPYREISNP 290  
 Db 222 LDYNSKNSBDDTNAKOKUKVLTJTSKASNSDDEHKNSRNSTMLMPYREISNP 277  
 Qy 291 IDISGKTBQLDKLDRKKAKANNDYNGVLOGAIIHGARETPKKEKKGPSGRHFLP 350  
 Db 278 IEL---TTKNDACKEWQDQAKDWDGVDLQGAIIHARBEITKEKEKSK---KROHTVLP 332  
 Qy 351 SODCRSTSYELONSVRBDTCKLSLGSAGTWSNLPWPPIFORTKHDQIDMDDVNUK 410  
 Db 333 SODESTFSYDHNK-SDSKLKLKURVNENITTSNLFPWLPFNPHTNRKADMIDYKLIK 391  
 Qy 411 LGQTGIAQDNIQSTSLISTSSLAGAFLGGSLSLTBLTKEYKSGDPLKENQDFYTER 470  
 Db 392 WGBKLIGIBGLDNLNTKLQAGASGIVGGPLGGSLSTBYLSKRYQSDRINASOFYER 451  
 Qy 471 VGECHVHFHSPEKKTGELPPKSETEPKKELPFENKNKNDKSTWEMFDKLSLTEROK 530  
 Db 452 VGECHVHFHSPEKKTGELPPKSETEPKKELPFENKNKNDKSTWEMFDKLSLTEROK 506  
 Qy 531 AKQSTIMKLUFLYTFKREHYTNYNLSATBRAKMAOEGITFYSVDTDLKTKRVEQ 590  
 Db 507 AKSEBAILKVLVYFFYKRDYVYHNNLSAIBRAKMAOEGITFYSVDTDLKTKRVEQ 566  
 Qy 591 VESTDKKKCREDTEKEHNEKUNTKOMSKBCKDPRDVDAKFKDILNTVUTET 650  
 Db 567 VKSERDKKKCREDTEKEHNEKUNTKOMSKBCKDPRDVDAKFKDILNTVUTET 626  
 Qy 651 PEDGVNKVNSQVSSENNNLSHSNYSVTHKAASDASWLY-SNKESLTWTISKQX 709  
 Db 627 FGGNTVSSESSWWSKTSIGNSNG--SSSRVSYTKGRGSGLSLSFGSTKESLTWTISKQX 684  
 Qy 710 BAPRKNSLTPPKYKQVNPQKULDKNKR----TKRDSTENKSVTKDQISATVNYK 763  
 Db 685 QAPERSEKPLTLPKQVNPQKULDKNKR----TKRDSTENKSVTKDQISATVNYK 744  
 Qy 764 INN-QPVKGMLDUDVLTYPETPTVPUVDPVSEHW--PIPEKPLVSEMPTRXPATNYPK 820  
 Db 745 INNGKOTNNNNLURERVNSYSPKMPBPEDLIEVVVPQVPELVLSPMLPPLPAIPIY-- 801  
 Qy 821 DIPTPOLPKDQDLESGGHGSVTDTGTGAEAGAQNQVSTOEMRPVIDETOP 880  
 Db 802 --PTPOLPKDQDLESGGHGSVTDTGTGAEAGAQNQVSTOEMRPVIDETOP 859  
 Qy 881 GMGSNDATVVEDTAKRPDVTLGGOSDPLDITEDQPSISGSNDATVVEDTVPKRD 940  
 Db 860 GMGSNDATVVEDTVPKRDQPSISGSNDATVVEDTVPKRD 919  
 Db 941 SLVGGSDPDTDTPGMSGSNGATWBTUTPKVEPHPNPQPEKNEOPSISLP 1000  
 Db 920 ILVGGSDPDTDTPGMSGSNDATVVEDTVPKRDQPSISGSNDATVVEDTVPKRD 979  
 Db 1001 QAPVVKAHLPASGDQKRASTIAAPPIIAGAGLISKRRTEGN 1046  
 Db 980 QAPVVKAHLPASGDQKRASTIAAPPIIAGAGLISKRRTEGN 1025

RESULT 2  
 S33850 fibronectin-binding protein - *Streptococcus dysgalactiae*  
 C;Species: *streptococcus dysgalactiae*  
 C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 09-Jul-2004  
 C;Accession: S33850; 832632  
 R:Lindren, P.B.; McGavin, M.J.; Sigmaeas, C.; Guss, B.; Gurusiddappa, S.; Hoeook, M.; Li, Eun-J. Blocham, 21, 819-827, 1993  
 Article: Two different genes coding for fibronectin-binding proteins from *Streptococcus dysgalactiae*, number: S33850; PMID:133307299; PMID:183196991

A;Molecule type: DNA  
 A;Residues: 1-1091 <LIN>  
 A;Cross-references: UNIPROT:Q06556; UNIPARC:UPI00000B92C0; EMBL:222150; NID:g288968; PI  
 Query Match 39.9%; Score 2138.5; DB 2; Length 1091;  
 Best Local Similarity 42.3%; Pred. No. 1.5e-80;  
 Matches 487; Conservative 184; Mismatches 310; Indels 171; Gaps 28;  
 Db 61 GAATTTTATTATGCGTOSTPABATPQPOAQIAAVATSTSSSDG----- 113  
 Qy 37 -ASVTTESPAIQTEEDQGQAAEELPTPAPOTSPSTSVAPARAJAMADEKGIAARAH 95  
 Db 114 -KAPQA--VTSSTSPSPAAASN-----GSQQAEEATEPOMEVEKYTDKEN 160  
 Db 96 EPAPKASYOAEASPGAAEATNTGOPTNTQOARSRSKRAAIAPQIIEVERLUDKEN 155  
 Db 221 QIDEGADYMALDQVKSKQSEDEPNNAKIKKUCLVKTJSKASNSDDEHKNSRN 280  
 Db 161 SKUNIKODCKTPKGSSVANEKOTKLIRURDGKLRDVTWTKTNEGTDIVTVTPK 220  
 Db 156 SSLTVQDG-----EOKOLIKRGHDQHDIRDTSRQVKGMDGTDVLTWTPK 204  
 Qy 205 OIDGEGAETVILDPTQSKTETDNTAKENIKLVTGTTGTDKGKNSH-YNARNSVRL 263  
 Db 281 MTFYRBRINPIDSGKTBQDLDKULDKRKKANTD-WGYDLOGAIIHAKABEINCEK 339  
 Db 264 IDPFRKVERSTDLGQMDAKKIDBKLNFWKKAQDQDYYNGWVULQGAIHARBEIT 323  
 Db 340 KFGKRRHIVLFSQAGESTSYELONSVREBDTCKLSLGSAGTWSNLPWPPIFNHFKNI 399  
 Db 324 RSGKROHTLPSQESTSYDPSKDKSDKNDQVAVE--PVTVSNPLFWPFPYDFTTRH 380  
 Qy 400 DMLDQVVKULVKLGCGTIGAGLDMQSTLSTLSSLAGAFLGGSLTEILILKVEKSD 459  
 Db 381 NVUNDAKCQKULDFLKNGKQSQFNGAVDNVATVNTGTLGSGPFLKNPQDYLISLADETSK 440  
 Qy 460 LKEQDVKYKVRGEGYHHSFSPERKKQKQCRIPKPKSEKPKKILPENNKNNDQKSWTENIF 519  
 Db 441 LNSEKPDYSSRVGQYHFRSFDR-EVKQVKPKLVLVEKIR--GMLKQPKQDWTI 495  
 Db 520 DKLSITERQKQAKOHTMLKLETFKQKBYHVNNTSIAEAKWMAOEGITFYSVDPID 579  
 Qy 496 SSSLGANSIREKIQDMMIDALDULPYRQYQOFYHNLSAQEARMARBERGKIQAVDNT 555  
 Db 580 LKTTKRRKQVES---TDDKKEKREDIEKERNENNDYKQMSGSKDPRFDDKA 635  
 Db 556 PERIAKEINSQKSYEAYTNHLKQGABAKBLAKGENKEDQYKLMGSBQK-FKDVEP 614  
 Qy 636 EKFKDITNTVUTEPEGDVNVKNSWYSSSENNNLSHSNYSVTH-KAASDASWLY- 693  
 Db 615 EKPKDITLTKEVTEPEKSV-----NNSQEQSKNEVYKQSSNSSLSPR 664  
 Qy 694 --SNKESLTWTISKQKRAFEPKNSLTFKQKQKULDKNKR-TKDST-ENKTS 750  
 Db 665 SSSTNISITWTLISKQKULQSGSTTLEYKLUHKDKPQKLAQTRKRSLSDTSENKKS 724  
 Qy 751 VTKDLSNTWVKINQSVKQKADWKLKVLTKEVTPVPUVGEVUPIPKLVEMTPL 810  
 Db 725 VTKVNTSVDVCKKINDKKEKCKEKLDSVLSVSKETTRPKQPTPV--RPTPKRPLTPL 781  
 Qy 811 YPAIPNP-----TPDIP-----TQQLRDEPLISGAGHPGSVTDIVTGTGARG-- 855  
 Db 782 APSEPSQSPSIPFPLPSRSPVPSFSPETPEGPGENL-GQSBRITTEQSOSENK 840  
 Qy 856 --GAQGVVSTOBENDPIV-----DIVEDTOPGMSGSNDATVVEDTAKRPDV 904  
 Db 841 PASGNETVVEDPQTSQDPSDIVVGGPQWIDFTEDSQGMSGNSHTEDSKPSQDDEVII 900  
 Qy 905 GQSDPDTDTPGMSGS-----DATVVEDTVPK----- 937

Pri or odd (2)

Page 1

Copyright (c) 1993 - 2006 Biocceleration Ltd.  
GenCore Version 5.1.9

30	219.5	4.5	2738	2
31	238.5	4.5	B88320	
32	237	4.4	1163	2
32	237	4.4	G97236	
33	237	4.4	A71523	
33	237	4.4	1192	1
			HMSQ1P	
34	235.5	4.4	1296	1
34	235.5	4.4	T28680	
35	236.5	4.4	1166	2
35	236.5	4.4	3488	2
36	236	4.4	T24418	
36	236	4.4	3523	2
37	235	4.4	T8477	
37	235	4.4	1381	1
38	234	4.4	S45781	
39	234	4.4	1822	
39	234	4.4	S33441	
40	233.5	4.4	T28676	
41	232.5	4.3	2401	2
41	232.5	4.3	S64339	
42	231.5	4.3	816	2
42	231.5	4.3	T27935	
43	231	4.3	1690	2
43	231	4.3	T13030	
44	231	4.3	1256	1
44	231	4.3	A41829	
45	230.5	4.3	1473	2
45	230.5	4.3	A35186	
45	230.5	4.3	T28688	
			protein F0A11.6	1
			ATPase involved in	
			probable secreted	
			aggregation protein	
			fibrinogen-binding	
			hypothetical prote	
			hypothetical prote	
			probable calcium-b	
			BP protein - StreB	
			rhopty protein -	
			hypothetical prote	
			hypothetical prote	
			microtubule bindin	
			muramidase-release	
			salivary agglutini	
			major merozoite su	

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

ALIGNMENTS

Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0% Maximum Match 100% Listing first 45 summaries  
Database : PIR 80:  
  
C.Species: *Streptococcus Pyogenes*  
C;Date: 14-Feb-1997 #Sequence\_revision 13-Mar-1997 #text\_change 09-Jul-2004  
C;Accession: S6790; S72666  
R.Krekemeyer, B.; Talay, S.R.; Chhatwal, G.S.  
Mol. Microbiol. 17, 137-145, 1995  
A;Title: Characterization of a novel fibronectin-binding surface protein in group A streptococci

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

No.	Score	Match Length	DB	ID	Description
1	3346.5	62.5	1025	2	fibronectin-bindin
2	2130.5	39.9	1091	2	fibronectin-bindin
3	409.5	7.6	659	1	adhesin - streptoo
4	394	7.4	638	2	fibronectin-bindin
5	354	6.6	961	2	hypothetical protein
6	351.5	6.6	1018	2	fibronectin-bindin
7	339	6.3	1038	2	hypothetical protein
8	338.5	6.3	940	2	fibronectin-bindin
9	325.5	6.1	1039	2	protein F2 - Sire
10	306	5.7	1109	2	fibronectin-bindin
11	302	5.6	1164	1	FC50AG
12	289	5.4	1134	2	IGA_Fc receptor p
13	282	5.3	1092	2	fibronectin-bindin
14	268.5	5.0	1875	2	myosin-like prote
15	263	4.9	1726	1	major merozoite s
16	263.5	4.9	1726	1	myosin heavy chain
17	263	4.9	2139	2	FmTB protein [imp
18	257.5	4.8	2481	2	hypothetical protein
19	248.5	4.6	2722	2	centromere proeli
20	248	4.6	2663	1	mature-parasite-1
21	247	4.6	1526	2	major merozoite s
22	245	4.6	1026	2	hypothetical protein
23	241.5	4.5	853	2	rhoptry protein -
24	241	4.5	2269	2	aggregation prote
25	240.5	4.5	1306	2	ORF MSV156 hypoth
26	240	4.5	1127	2	hypothetical protein
27	240	4.5	1141	2	membrane nucleosae
28	239.5	4.5	1125	2	surface-associatee
29	239.5	4.5	2508	2	S61441

Description	Start codon:	End codon:	Length:	Sequence
A; Molecule type: DNA	A;	T;	1-19;	'T'.
A; Cross-references: UNIPARC:UPI00000AF5BB; EMBL:X83303; NID:91070387; PIDN:CAAS58282.1;				<KRW>
C; Genetics:				
A; Gene: sfbII				
A; Start codon: TAG				
C; Keywords: fibronectin binding; membrane bound				
F; 1-34/Domain: signal sequence #status predicted <SG>				
F; 35-1025/Product: fibronectin-binding protein II #status predicted <MAT>				
F; 848-868/Region: fibronectin binding				
protein F2 - Strept				
Fibronectin-bindin				
IgA Fc receptor pr				
IgA Fc receptor pr				
fibronectin-bindin				
myosin-like protei				
major merozoite su				
merozoite su				
myosin heavy chain				
Fmcb protein limpo				
hypothetical prote				
centromere protein				
major-parasite-in				
hypothetical prote				
rioptery protein -				
aggregation prote				
Our MSL15 hypothe				
hypothetical prote				
membrane nuclease				
surface-associated				

**A:**Molecule type: DNA  
**A:**Residue: 1-1091 <LIN>  
**A:**Cross-references: UNIPROT:Q06556; UNIPARC:UPI00000B92C0; ENBL:Z22150; NID:928968; P11



Db 61 GAAVTATTATNGQPSTPVARATPQPQQTAPVALTSSASSSDGKAPQAVT 120  
 QY 121 SSTSPTPAASSNSNSNORASARTBQTMEYKTYTDKENSINTKGKPKTSSWNR 180  
 Db 121 SSTSPTPAASSNSNSNORASARTBQTMEYKTYTDKENSINTKGKPKTSSWNR 180  
 Qy 181 KDTKLIRNROCKLARDIVRTVKINEDGIDNTVVKPKQIDEADWMLDVSKRSE 240  
 Db 181 KDTKLIRNROCKLARDIVRTVKINEDGIDNTVVKPKQIDEADWMLDVSKRSE 240  
 Qy 241 DDFNAKRNKKLVKLTISGASNSNDREHKNRSNSVRALTFYRIBISNPIDIGKTEQ 300  
 Db 241 DDFNAKRNKKLVKLTISGASNSNDREHKNRSNSVRALTFYRIBISNPIDIGKTEQ 300  
 Qy 301 LDKLUDRLRKAKANVDGVDLQAGAIHKARIBIENKEKEKGKGRHIVPSQESTSR 360  
 Db 301 LDKLUDRLRKAKANVDGVDLQAGAIHKARIBIENKEKEKGKGRHIVPSQESTSR 360  
 Qy 361 LONSUREDKWLSLSGATVSSNPLPPWPPIPNTHKNDLDDYKNUKLGONGIAQL 420  
 Db 361 LONSUREDKWLSLSGATVSSNPLPPWPPIPNTHKNDLDDYKNUKLGONGIAQL 420  
 Qy 421 DNQSTLSLISTGSSLAGTAGGGSLETVILKEWKGDIKENPDYTGVGEHMFHFP 480  
 Db 421 DNQSTLSLISTGSSLAGTAGGGSLETVILKEWKGDIKENPDYTGVGEHMFHFP 480  
 Qy 481 SERKKGTEPPKSIEBKPKCIELFENKNDKWSWTEWPKSLTERICKAOETMIL 540  
 Db 481 SERKKGTEPPKSIEBKPKCIELFENKNDKWSWTEWPKSLTERICKAOETMIL 540  
 Qy 541 BYLFYKRETYNHNLSAEEAKMAQEGITFYSVDTLKTTSKRVKQVESTDKCR 600  
 Db 541 BYLFYKRETYNHNLSAEEAKMAQEGITFYSVDTLKTTSKRVKQVESTDKCR 600  
 Dr 601 KOREDIKEKNEKPNYKOMSECKDPRFDVKAELKFOLINTVTTFEDQWVKD 660  
 Db 601 KOREDIKEKNEKPNYKOMSECKDPRFDVKAELKFOLINTVTTFEDQWVKD 660  
 Qy 661 SWQISSENNNLHSNYSKTHKAASDASHWLYSNKESTJTWISKQOLCAFEKNSJTF 720  
 Db 661 SWQISSENNNLHSNYSKTHKAASDASHWLYSNKESTJTWISKQOLCAFEKNSJTF 720  
 Qy 721 KYLGQVNKOKLDKNKRTRDTSTENKSVTKOISNTWYKINNOETKGKNUDDVKT 780  
 Db 721 KYLGQVNKOKLDKNKRTRDTSTENKSVTKOISNTWYKINNOETKGKNUDDVKT 780  
 Qy 781 YTKEETVPUVDEVEGYVPIERKPLTERMTLYPALPNYPDPIDPOLPQDKEBASGGG 840  
 Db 781 YTKEETVPUVDEVEGYVPIERKPLTERMTLYPALPNYPDPIDPOLPQDKEBASGGG 840  
 Qy 841 PSYDIVPDTGAGGAQNGWVSOENRPVIDTEDTPQGMSSNDATVVEDTAPKRP 900  
 Db 841 PSYDIVPDTGAGGAQNGWVSOENRPVIDTEDTPQGMSSNDATVVEDTAPKRP 900  
 Qy 901 DVLGQQSDPIDIETDQSVGSSNDATVVEDTAPKRDSDLVQGQSDPIDIETDQGW 960  
 Db 901 DVLGQQSDPIDIETDQSVGSSNDATVVEDTAPKRDSDLVQGQSDPIDIETDQGW 960  
 Qy 961 SGSGNATVTFEDTRPKRVHPDNFOPAPKPNBOPSLSPQAPYKAMHLPASGDKERA 1020  
 Db 961 SGSGNATVTFEDTRPKRVHPDNFOPAPKPNBOPSLSPQAPYKAMHLPASGDKERA 1020  
 Qy 1021 SFTIAAPTIGAAGLISKRDRTEGN 1046  
 Db 1021 SFTIAAPTIGAAGLISKRDRTEGN 1046

RESULT 2  
 Q93751 STRPY PRELIMINARY; PRT; 1027 AA.  
 ID Q93751-STRPY PRELIMINARY; PRT; 1027 AA.

6/6/2006 6:13:33 2006

07-PFB-2006, entry version 16.  
 DT Serum opacity factor Vn21.  
 DB Streptococcus pyogenes.  
 OS Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;  
 OC Streptococcus.  
 OX NCBI\_TaxId=1314;  
 RN [1]  
 RP NUCBOTTDE SEQUENCE.  
 RX MEDLINE-2164287; PubMed=11782509;  
 RA Gilman C.M., Towers R.J., McMillan D.J., Delvecchio A.'s.,  
 RA Sriprakash K.S., Currie B., Kreikemeyer B., Chhatwal G.S.,  
 RA Walker M.J., "Immunological response mounted by Aboriginal Australians living in  
 RT the Northern Territory of Australia against Streptococcus pyogenes  
 RT serum opacity factor.", Microbiology 148:169-178 (2002).  
 RL  
 CC Copyrighted by the UniProt Consortium, see <http://www.uniprot.org/terms>  
 CC Distributed under the Creative Commons Attribution NonDerivs License  
 CC  
 DR EMLC; ARX367014; ANX5268\_1; Genomic\_DNA.  
 DR GO; GO:000986; C:membrane; IEA.  
 DR GO; GO:00116020; C:membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR004237; Fibron repeat\_bd.  
 DR InterPro; IPR005877; Gpos\_VSIRK.  
 DR InterPro; IPR01999; Gram\_Pos\_anchor.  
 DR InterPro; IPR020235; VWF\_A.  
 DR Pfam; PF02986; Fn\_bind; 3.  
 DR Pfam; PF00746; Gram\_pos\_anchor; 1.  
 DR Pfam; PF00052; VWA; 1.  
 DR SMART; SM00327; VWA; 1.  
 DR TIGRFAMs; TIGR01168; VSIRK\_signal; 1.  
 DR PROSITE; PSS0847; GRAM\_POS\_ANCHORING; 1.  
 DR PROSITE; PS00234; VWF\_A; 1.  
 SQ SEQUENCE: 1027 AA; 112542 MW; 04823958598ACTA CRC64;  
 Qry Best Local Similarity 66.7%; Score 3575.5; DB 2; Length 1027;  
 Matches 751; Conservative 98; Mismatches 153; Indels 69; Gaps 26;  
 Qy 1 MTNCKYKURKLSVGLSVGTMLIAPTYLQEVAST--TSTETSTASAGYTGTAASE 57  
 Db 1 MTNCKYKURKLSVGLSVGTMLIAPTYLQEVAST--TSTETSTASAGYTGTAASE 57  
 Qy 58 TGGAAVTATT-----ATINGGPOSTPVARATPQQTAPVAAT 102  
 Db 55 TTSQTATTTPSAITDAGCAGSGTNQGSSVTSQGSS-GTTTAPRPOPOATT-AAT 111  
 Qy 103 STSSASSSSDGKARQAVTSSTSPPTAASSNSNSNORASAEATEPQTMVEVKYTDKNSK 162  
 Db 112 SISS-NSSSDGQ1PK--TAITPSTPAAASNN-SNOBASAEATEPQTMVEVKYTDKNSK 166  
 Qy 163 LNTRKDGPKPKTSSVNNEDTKLIRNEDGKARDIVRTVKINEDGIDNTVVKPKQI 222  
 Db 167 AKYKDGPKNSSV--DDKDTKLIRNEDGKARDIVRTVKINEDGIDNTVVKPKQI 225  
 Qy 223 DEGRDWMLDVSKOMSBDDEFFNAKOKKLYKLTISKSASNSNDHEKMSNSRFLMT 282  
 Db 226 DEGRDWMLDVSKOMSBDDEFFNAKOKKLYKLTISKSASNSNDHEKMSNSRFLMT 279  
 Qy 283 PYRISUPIDISGKTEQDLDKLDLRRKAKANVDGVDLQAGAIHKARIBIENKEKGK 342  
 Db 280 PYRISUPIDISGKTEQDLDKLDLRRKAKANVDGVDLQAGAIHKARIBIENKEKGK 339  
 Qy 343 KRKHIVPSQGSTFSVYELONSVREDKWLSLSGAVTSSNPLPPWPPIPNTHKNDML 402  
 Db 340 KHOHIVPSQGSTFSVYELONSVREDKWLSLSGAVTSSNPLPPWPPIPNTHKNDML 396  
 Qy 403 DDVKNLUKGQGTTGIGLQDNTQSLISTSSLAGAFLAGSGSLTUYLTGKSYGDLK 462  
 Db 397 DDKIKLWKWGRAGIGBGLGSKDQKUSLGLASSAAGLFOEGSLTUYLSLKYSOSLSKA 456

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- BLACK BORDERS**
- IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- FADED TEXT OR DRAWING**
- BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- SKEWED/SLANTED IMAGES**
- COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- GRAY SCALE DOCUMENTS**
- LINES OR MARKS ON ORIGINAL DOCUMENT**
- REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- OTHER:** \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**